Creation and validation of tool to assess resident competence in neonatal resuscitation.

Sara K. Kane, DO and Diane E. Lorant, MD
Indiana University School of Medicine
699 Riley Hospital Drive, RR208
Indianapolis, IN 46202

(317) 274-4715

skkane@iu.edu
dlorant@iu.edu

Corresponding author: Sara K. Kane, DO

Acknowledgements: The authors would like to thank the simulation experts at the Indiana University Simulation Center for their expertise and participation in the project. Statistical analysis was completed by George J. Eckert, MS, Indiana University.

This project was funded with support from the Indiana Clinical and Translational Sciences Institute funded, in part by Grant Number UL1TR001108 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. The sponsor had no role in the study design, collection, analysis or interpretation of data.
There are no conflicts of interests or corporate sponsors for this project.

ABSTRACT

Background: The American Board of Pediatrics requires that pediatricians be able to initiate stabilization of a newborn. After residency, 45% of general pediatricians routinely attend deliveries. However, there is no standard approach or tool to measure resident proficiency in newborn resuscitation across training programs. In a national survey, we found a large variability in faculty assessment of the amount of supervision trainees need for various resuscitation scenarios. Objective documentation of trainee performance would permit competency-based decisions on the level of supervision required and facilitate feedback on trainee performance.

Methods: A simplified tool was created following the Neonatal Resuscitation Program (NRP) algorithm, with emphasis on communication, leadership, knowledge of equipment, and initial

stabilization. To achieve content validity, the tool was evaluated by the NRP steering committee. To assess internal structure of the tool, we filmed 10 simulated resuscitation scenarios 9 of which contained errors. Experienced resuscitation team members used the tool to assess performance of the team leader in the videos. To evaluate the response process, the tool was used to assess experienced resuscitators in real time at academic and non-academic sites.

Results: The NRP steering committee approved the tool, providing evidence of content validity. Performance of the team leader in the simulated videos was assessed by 16 evaluators using the tool. There was an intra-class coefficient of 0.86 showing excellent agreement. There was no statistical difference in scores between 102 resuscitations led by experienced resuscitators at academic and non-academic hospitals (p=0.98), which demonstrates generalizability.

Conclusions: The tool we have developed to assess performance in initiating newborn resuscitation shows evidence of construct validity based on assessment of content and internal structure (inter-observer agreement, response processes, and generalizability).

Keywords: Neonatal resuscitation; EPA; resident competence

What's New: We developed and provided validity evidence for a tool that can be used by members of the resuscitation team to efficiently assess the performance of trainees leading a resuscitation. The tool has potential to facilitate education in newborn resuscitation by providing immediate feedback to trainees on their performance.

MANUSCRIPT TEXT

Background

The Accreditation Council for Graduate Medical Education (ACGME) requires training programs to provide data to medical personnel regarding the level of supervision each trainee should have for specific clinical scenarios, including delivery room resuscitation of a newborn. The traditional model has been that after successful completion of the Neonatal Resuscitation Program (NRP) and one year of post-medical school training, residents are entrusted to lead uncomplicated neonatal resuscitations without supervision or formal assessment of their skills.[1] However, studies have shown that competence drops off with time after completing an NRP course.[2] We surveyed neonatologists across the country and found that the amount of supervision trainees receive in the delivery room is inconsistent and often dependent on the postgraduate year of training.[3] There is no standard approach or tool to measure performance in newborn resuscitation outside of simulation.

The American Board of Pediatrics (ABP) expects that pediatricians should be able to initiate stabilization of a newborn, as evidenced by the Entrustable Professional Activity (EPA) "provide resuscitation and stabilization of neonates and infants that aligns care with the severity of illness".[4] A survey of general pediatricians indicates that 45% routinely attend deliveries after residency,[5] making an objective assessment of trainee performance in neonatal resuscitation an important component of their training.

Given the limited number of neonatal resuscitations trainees attend, postgraduate year of training is no longer an adequate surrogate for competency. The lack of longitudinal assessment of individual trainees makes it difficult for faculty to assess their ability to resuscitate a newborn. Objective assessment of trainees' skills would allow faculty to make competency-based decisions on the level of supervision necessary. The aim of this study was to create and provide validity evidence for a novel tool to assess trainee performance of neonatal resuscitation.

Methods

The 7th edition of the Neonatal Resuscitation guidelines up to, but not including, intubation of the neonate, was used as a blueprint for the tool. Emphasis was placed on communication, leadership, knowledge of equipment, and initial stabilization. Competence in intubation was not included because the updated neonatal resuscitation guidelines no longer recommend routine intubation for non-vigorous meconium deliveries.[6]

The tool was designed so that it could be completed quickly and in real time. It consists of ten yes/no questions that can be filled out by the evaluator during or immediately after the newborn resuscitation (Figure 1).

The first step in providing validity evidence for the tool was to examine the content. Since content domain is usually left to the judgement of subject matter experts,[7] the tool was sent to the NRP Steering Committee of the AAP for review and approval.

The next step in providing validity evidence was to assess the internal structure of the tool by measuring inter-observer reliability. To achieve this, we developed ten scenarios in which a trainee would be expected to initiate neonatal resuscitation. One of the ten scenarios had no errors, and the other nine contained one or more common errors frequently performed by

trainees. Scripts were written by the primary investigator and were enacted by trained simulation instructors using a high-fidelity newborn mannequin (SimNewB). The scenarios were videotaped at the Indiana University Simulation Center with three cameras positioned to record different angles of the resuscitation. In order to determine if individuals from various disciplines had a similar assessment of the resuscitation, the videos were assessed by four neonatologists, four neonatal fellows, four neonatal nurse practitioners, and four respiratory therapists. Individuals were trained in using the tool by the primary investigator, and were given copies of the tool with instructions on its use. Individuals were asked to watch each scenario once without rewinding and use the tool to score the trainee's simulated performance. The scores were compared within each discipline and between the entire cohort of evaluators to assess for inter-observer reliability, calculated by intraclass correlation coefficients.

Before using the tool to evaluate trainees in newborn resuscitation, we needed to ensure that experienced resuscitators usually performed each item on the checklist correctly. The tool was used to evaluate resuscitations performed by neonatologists, neonatal fellows, and nurse practitioners at one academic institution and three community hospitals that have high volume delivery services. As in the simulated scenarios, team members performing the assessment were trained by the primary investigator, and were given copies of the tool with instructions. We then evaluated the performance of these skilled resuscitators between disciplines and compared results between the academic and non-academic institutions. Logistic regression was used to compare performance between groups.

Approval for this study was obtained from hospital sites' Institutional Review Boards (IRB). The IRBs did not require informed consent for this study.

Results

The NRP Steering Committee reviewed and approved the tool, which provided validation evidence of the tool's content. Their approval was contingent upon addition of the third item on the checklist, "Does the resident address delayed cord clamping?" which is recommended in the 7th edition of the Neonatal Resuscitation guidelines.[6]

Inter-observer reliability was measured using the videotaped simulation scenarios and comparing the results of the evaluators' score sheets with the known correct and incorrect maneuvers during the simulated scenario (Table 1). The percent of evaluators who correctly identified errors was 97%, and there was no significant difference between the four groups of evaluators (neonatologists, neonatology fellows, neonatal nurse practitioners, and respiratory therapists, p=0.77). Question 5 (titrating FiO2 based on minute of life guidelines) was the least likely to be identified correctly by all disciplines evaluating the videos. To determine if each discipline was equally adept at using the tool, we calculated intraclass coefficients for each question on the checklist across all groups of observers (Table 2). The overall intraclass coefficient was 0.86, showing excellent agreement between all observers. Question 5 had the lowest intraclass coefficient of 0.45, which shows only moderate agreement. When that question was removed from analysis, the overall intraclass coefficient increased to 0.90.

To assess the tool's response process and generalizability, we used the tool to evaluate experienced practitioners at one academic and three community sites. We examined the percentage of checklist items that were performed correctly by the neonatal resuscitator, and compared academic and nonacademic sites. There were 62 observations performed at the academic center and 40 performed at community hospitals. The checklist indicated that the entire resuscitation was performed correctly 90% of the time at the academic site, and 50% of the time at community hospitals (p=0.0006). Upon examining the data, Question 3 (assessing delayed

cord clamping) was completed 77% of the time among all sites, with a large difference between the academic and community sites (90% vs. 37%). With removal of Question 3 from analysis, the checklist indicated that resuscitation was performed correctly 100% of the time at the academic hospital and 83% of the time at the community sites. There was no statistical difference between the academic vs. community sites (p=0.98). There were no adverse events during the observed resuscitations as a result of using the tool for assessment.

Discussion

The ACGME is moving towards EPAs as a more formal and objective way of evaluating residents. An EPA is a "task or responsibility that can be entrusted to a trainee once sufficient competence is reached to allow for unsupervised practice" [8]. The pediatric EPA "resuscitate, initiate stabilization and triage to align care with severity of illness", requires multiple complex competencies that would be utilized in the resuscitation of a neonate. Although few neonates require full resuscitation after delivery, 10% need some assistance to establish ventilation.[6] While in situ simulation training has facilitated education in neonatal resuscitation, performance following simulation is not necessarily transferable to the clinical environment. [9, 10] We previously studied the amount of supervision neonatologists thought necessary for neonatal resuscitation, and found significant variability.[3] Therefore, an objective assessment of trainees' ability to lead a neonatal resuscitation is needed.

We developed an objective tool to assess trainees' performance in neonatal resuscitation and provide evidence of construct validity, specifically in the areas of content and internal structure (response processes, inter-observer reliability, and generalizability).[11, 12] Content validity can be defined as the degree to which elements of an assessment instrument are relevant

to and representative of the targeted construct for a particular assessment purpose.[13] Provider knowledge, skills, and performance of NRP have been shown to support the successful transition at birth and reduce infant mortality.[14] Therefore, we sought to establish content validity by the NRP Steering Committee approval.

Evidence of the tool's internal consistency was achieved by using the videotaped simulation scenarios and comparing the results of the observers' score sheets with the known correct and incorrect maneuvers during the simulated scenario. There was no statistical difference in the evaluation of the videos between disciplines (neonatologists, neonatology fellows, neonatal nurse practitioners, and respiratory therapists, p= 0.77). Respiratory therapists are an integral component of pediatric resuscitation teams.[15] Their ability to use the tool to assess trainee performance in neonatal resuscitation provides the opportunity to evaluate every resuscitation a trainee leads. This makes the tool applicable to real-life situations, in which additional team members may not be available solely to evaluate a trainee's performance.

Although the overall inter-observer reliability showed excellent agreement, Question 5 on oxygen administration per minute of life guidelines had the most variability among observers. Upon further review of our simulated videos, the minute of life timer was not easily visible to the evaluator. Because this was a problem with the video and not the evaluator, we analyzed the reliability both with and without Question 5. Without Question 5, the intra-class coefficient showed near perfect agreement. However, titrating oxygen based on minute of life guidelines may be challenging in real-time resuscitations. Video recordings from the delivery room of tertiary care hospitals show that many infants don't have resuscitation tasks completed within the time frame recommended by NRP.[16]

The response process and generalizability were assessed by using the tool on experienced practitioners at one academic and three community sites. There was significant discrepancy of checklist performance between academic and non-academic sites (90% vs. 50%, p=0.0006). Question 3 (assessing delayed cord clamping) was the major contributor, with 90% vs. 37% completion between the academic and non-academic sites respectively. One of the community sites had not yet adopted delayed cord clamping because this is a new addition to the most recent edition of the NRP guidelines. After removing this question, there was no difference in performance between the academic and non-academic sites (p=0.98). Adding cord clamping to the tool emphasizes the benefits of delayed cord clamping and is an important educational aspect of the tool.[17]

We created and provided validity evidence for this tool to provide neonatologists, neonatal fellows, and neonatal nurse practitioners an objective way to determine the level of supervision a trainee may need to lead a neonatal resuscitation. The tool is short, easy to use, and can be completed in real time. It has universal adaptability to academic and community delivery centers, and doesn't require additional personnel to be present for the evaluation of the trainee. However, there are several limitations to the tool's utility. Since most infants require minimal resuscitation, a trainee may appear competent to lead a resuscitation unsupervised by performing the initial steps in resuscitation correctly but not be able to successfully complete all the steps up to intubation. The assessment tool does not include intubation because delivery room intubation is less frequent now that it is no longer recommended to intubate and suction a non-vigorous infant with meconium stained fluid.[5] However, intubation may be unanticipated and is a required skill by the ABP for all pediatricians.

Although experienced resuscitators performed the entire checklist correctly 90% of the time, we have yet to establish a passing score for trainees. An additional study is ongoing to determine how many successful resuscitations a trainee needs to lead before consistently performing each step in the tool correctly. This study is designed to evaluate the intended and unintended consequences of the assessment tool before using it to attest competence. In the meantime, the tool is used to educate trainees and give immediate objective feedback to refine their resuscitation skills.

The authors would like to thank the members of the Indiana University Simulation Center for use of their facility and equipment and George Eckert, MS for statistical analysis. Indiana University holds the rights to this tool. This project was funded with support from the Indiana Clinical and Translational Sciences Institute, by Grant Number UL1TR001108 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award.

References

1. Halamek LP, Kaegi DM. Who's teaching Neonatal resuscitation to housestaff?: Results of a national survey. Pediatrics 2001;107:249-55.
2. Patel J, Posencheg M, Ades, A. Proficiency and retention of neonatal resuscitation skills by pediatric residents. Pediatrics 2012; 3: 515-521.
3. Kane S, Lorant D. The amount of supervision trainees receive during neonatal resuscitation is variable and often dependent on subjective criteria. Journal of Perinatology 2018; 1-6.
4. Entrustable Professional Activities for General Pediatricians. 2013. (Accessed March 21, 2017, at https://www.abp.org/entrustable-professional-activities-epas.)
5. Walton DM, Edwards MC. Nationwide survey of pediatric residency training in newborn medicine: preparation for primary care practice. Pediatrics 2002;110:1081-7.
6. Wyckoff MH, Aziz K, Escobedo MB, et al. Part 13: neonatal resuscitation. Circulation 2015;132:S543-S60.

7. Lawshe CH. A quantitative approach to content validity. Personnel psychology 1975;28:563-75.
8 El-Haddad C, Damodaran A, McNeil HP, Hu W. The ABCs of entrustable professional activities: an overview of 'entrustable professional activities' in medical education. Internal medicine journal 2016;46:1006-10.
9. Rubio-Gurung S, Putet G, Touzet S, Gauthier-Moulinier H, Jordan I, Beissel A, et al. In situ simulation training for neonatal resuscitation: an RCT. Pediatrics 2014;134:e790–7
10. Finan E, Bismilla Z, Campbell C, Leblanc V, Jefferies A, Whyte HE. Improved procedural performance following a simulation training session may not be transferable to the clinical environment. Journal of Perinatology 2012;32:539–44.
11. Downing SM. Validity: On the meaningful interpretation of assessment data. Medical Education 2003; 37:830-7.
12. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: Theory and application. The American journal of medicine 2006; 119:166.e7-166.e16.
13. Haynes SN, Richard D, Kubany ES. Content validity in psychological assessment: A functional approach to concepts and methods. Psychological assessment 1995;7:238.
14. Niermeyer S. From the neonatal resuscitation program to helping babies breath: Global impact of educational programs in neonatal resuscitation. Seminars in fetal and neonatal medicine 2015; 20: 300-8.
15. Sawyer T, Laubach VA, Hudak J, Yamamura K, Pocrnich A. Improvements in teamwork during neoantal resuscitation after interprofessional TeamSTEPPS training. Neonatal network 2013; 1: 26-33.
16. McCarthy LK, Morley CJ, Davis PG, Kamlin COF, O'Donnell CP. Timing of interventions in the delivery room: does reality compare with neonatal resuscitation guidelines? The Journal of pediatrics 2013;163:1553-7. e1.
17. Delayed umbilical cord clamping after birth: American Academy of Pediatrics statement of endorsement. Pediatrics 2017; 139(6).

Figure 1: Neonatal Resuscitation Evaluation Tool	
1. Were team member roles and responsibilities clearly defined at the start of the resuscitation?	[ ] Yes [ ] No
2. Was resuscitation equipment checked?	[ ] Yes [ ] No
3. Does the resident address delayed cord clamping?	[ ] Yes [ ] No
4. Were the first steps of resuscitation (warm, dry, stimulate) accomplished?	[ ] Yes [ ] No
5. Did resident direct supplemental O2 administration per minute of life guidelines?	[ ] Yes [ ] No
6. Was mouth/nose suctioned before initiation of PPV?	[ ] Yes [ ] No
7. Does resident assess breathing after initial steps and initiate PPV within the first 60 seconds if apneic? If breathing, does resident evaluate HR and initiate PPV for HR <100?	[ ] Yes [ ] No
8. Does resident correctly initiate MRSOPA if HR does not improve?	[ ] Yes [ ] No
9. Does resident call for help if needed?	[ ] Yes [ ] No
10. Did the resident lead the resuscitation?	[ ] Yes [ ] No

Table 1: Inter-Observer Reliability, Percentage Correct by Groups of Evaluators					
Question	Overall	Faculty	Fellow	Nurse Practitioner	Respiratory Therapist
All	97%	97%	97%	96%	97%
1	99%	100%	100%	100%	95%
2	98%	98%	100%	95%	100%
3	99%	100%	100%	98%	100%
4	98%	98%	100%	98%	98%
5	84%	78%	90%	90%	78%
6	99%	100%	98%	98%	100%
7	97%	98%	98%	93%	100%
8	96%	100%	90%	93%	100%
9	99%	98%	100%	98%	100%
10	98%	98%	98%	100%	98%

Table 2: Inter-Observer Reliability, Intraclass Correlation Coefficients (ICCs) by Groups of Evaluators					
Question	Overall	Faculty	Fellow	Nurse Practitioner	Respiratory Therapist
All	0.86	0.85	0.88	0.84	0.91
1	0.94	1.00	1.00	1.00	0.76
2	0.90	0.87	1.00	0.76	1.00
3	0.97	1.00	1.00	0.90	1.00
4	0.86	0.79	1.00	0.79	0.79
5	0.45	0.20	0.56	0.63	0.70
6	0.95	1.00	0.90	0.90	1.00
7	0.84	0.84	0.84	0.69	1.00
8	0.85	1.00	0.73	0.76	1.00
9	0.94	0.87	1.00	0.87	1.00
10	0.83	0.79	0.67	1.00	0.79